

Implications of imposed nationwide testing (for accountability) on primary student learning

Nicole Mehrstens

College of Education, Health and Human Development, University of Canterbury, New Zealand

Abstract

Many education systems are increasing accountability measures through imposing standardised assessment practices. Assessment can have a significant impact on a young person's development therefore it is essential to explore implications of assessment for accountability. This review summarises key themes in contemporary literature regarding the implications of high-stakes imposed testing for accountability at primary level. The level of accountability within the New Zealand education model is then discussed.

Keywords: *standardised assessment, accountability, primary education, implications*



Journal of Initial Teacher Inquiry by [University of Canterbury](#) is licensed under a [Creative Commons Attribution 4.0 International License](#).

Permanent Link: <http://hdl.handle.net/10092/11440>

Assessment for Accountability

Many countries impose standardised academic testing in order to measure how well students have achieved. These test scores are believed to reflect ability in a domain of knowledge (Herman, 2005). Nationwide tests are often part of a regime to enhance 'accountability' in education, with the intended purpose of improving performance by examining its impact, measuring quality and results, and impelling staff to achieve higher standards (Garipey, Spencer, & Couture, 2009). Classroom assessment is considered to have a significant impact on a young person's academic life and development (Crooks, 1988; Herman, 2005), hence it is important to consider the implications of assessment and accountability measures for students. For purpose of simplicity in this review, 'assessment' refers to imposed, standardised, summative tests. The following piece is a summary of the key themes in, and implications of, research regarding widespread imposed testing on primary level student learning.

A brief exploration of the influences driving the increase of accountability in education is covered first. Two key themes from contemporary literature regarding the effects of accountability on primary level student learning are then summarised; the negative effects; and issues with the reliability of test scores. Finally, the level of accountability in the New Zealand education system is summarised and future implications are considered.

The Accountability era

Many education systems in worldwide are increasing accountability measures through imposed assessment practices. Sadler (2000) and the Ministry of Education (1994) assert that assessment can serve three purposes; to improve learning capability; for reporting; and for accountability for providing summative information. Some academics claim that accountability is predominantly about the effectiveness and 'value for money' of the education system (Knight, 2000; Sadler, 2000). Numerous academics assert that the ascendancy of managerial values and neoliberal governmental ideology are influencing the modern accountability era in education – an arena that was previously based on trust and professionalism (Herman, 2005; Knight, 2000).

Today, education systems in many countries are monitored and audited much like any other service – via imposed testing. Information gained on student achievement and progress is regularly reported to parents and summarised for public knowledge in the media (Knight, 2000); accountability has increased on multiple levels. Some education systems aim to enhance teacher accountability by adopting performance-related-pay programs, resource-allocation, sanctions and consequences, to raise the stakes according to students' test results (Garipey, Spencer, & Couture, 2009; Podgursky & Springer, 2007). Pilot-programs and research have highlighted inherent difficulties in reliably identifying effective teaching –

something that many academics argue cannot be measured accurately as it is influenced by multiple factors outside of the teachers' control (Podgursky & Springer, 2007).

Consequences associated with performance-related-pay measures include a reduction in teacher collaboration, and a shift in the view and nature of teaching and learning (Podgursky & Springer, 2007). Critics of the neoliberal and managerial approach of reducing education to simple 'measurable' outcomes, argue that this approach is impelling society to measure what can easily be measured, rather than what society values – thus impacting on the very nature and purpose of education (Biesta, 2010; Herman, 2005). Also, it is suggested by Knight (2000) that people who are systematically not trusted will eventually become untrustworthy. Assessment for educational accountability is influenced by neoliberal ideology and managerialism. This can lead to education being treated as a commodity.

Negative effects on student learning

A key theme in educational accountability literature is the negative effects of imposed testing on student learning. Research has acknowledged that test anxiety and assessment results can have a significant impact on a young person's development (Crooks, 1988; Herman, 2005; Marlow et al., 2014; Putwain, 2008; Shutz & Pekrun, 2011). In addition, researchers attest that imposed accountability testing puts significant pressure on teachers and schools and elicits higher stakes for outcomes (Crooks, 1988; Herman, 2005; Putwain, 2008). The literature strongly implies that this pressure has resulted in prioritised attempts to improve student outcomes in tested domains of learning, while other domains are neglected (Herman, 2005). This is referred to as the 'tested curriculum', signifying a distortion of the curriculum and standards (Herman, 2005). Many researchers consider such changes in teachers' use of classroom time concerning, as they were not motivated by any logical sense of curriculum or learning need (Herman, 2005).

Research also strongly suggests that educators who are under pressure to show improvement are 'teaching to the test'; focusing on test-relevant skills and formats, particularly in poor schools (Herman, 2005). It is widely believed that this results in students efficiently answering questions about prescribed portions of knowledge, while broader and deeper cognitive processes are not being intentionally developed, measured or valued (Herman, 2005; Knight, 2000). It is argued that such trends can demotivate students by impacting their confidence and ability to learn whilst students who work in narrowly constrained ways are rewarded (Knight, 2000; Marlow et al., 2012).

Another key theme in the literature is the importance of relationships, students' sense of connection, commitment to schooling, safety, positive norms, and efficacy, which are essential for effective learning and academic progress (Herman, 2005; Shutz & Pekrun, 2011). Many researchers claim that accountability measures cause a reduction in meaningful learning experiences and negatively impact academic outcomes (Herman, 2005). Thus many academics advocate the need for a balance of social and academic capital to be developed (Herman, 2005).

Unreliability of large-scale test scores

An additional theme highlighted within the educational accountability literature is the unreliability of large-scale test scores as reflective measures of student ability. Along with many others, Herman (2005) and Crooks (1988) state that all measures of student performance contain error therefore important decisions should never be based on a single test result (Crooks, 1988). In addition, Herman asserts that a test can only measure a portion of what students are learning and is therefore imperfect, and that it is impossible for tests to assess everything that is important.

Another aspect of dubious test score reliability is the restricted types of items that are included in tests according to their easiness to base questions on, whether they relate in empirically coherent ways with other items, and their level of difficulty (Herman, 2005). Research has shown that test scores typically increase substantially in the first three years of a new test being imposed, followed by a levelling-out (Herman, 2005). In addition, significant discrepancies between student performance on nation-imposed tests and other achievement measures with less substantial consequences, are commonly found (Herman, 2005). It is widely insinuated that this could indicate inflation of test scores, which would cause major discrepancies across the system (Brown, 2004; Herman, 2005).

In addition to these wider influences, there are multiple factors that affect individual students' (and schools') performances on accountability tests. For example, how they are feeling on the day, their carefulness in completing answer-sheets, attentiveness, proportion of high or low achievers away that day, reading and language levels, disabilities and learning styles, and so on (Herman, 2005). Thus a discrepancy between the intended aim of assessment and what it is actually quantifying, is commonly recognised (Marlow et al., 2012). Students are considered to be diverse in their styles of learning, therefore many academics question whether they should all be tested in the same way to avoid eliciting unreliable information (Knight, 2000).

The New Zealand Context

The New Zealand education model of evaluation and assessment is considered unique and is characterised by a high level of trust in schools and professionals (Crooks, 2011; Nusche et al., 2012). Each school is responsible for interpreting the New Zealand Curriculum and is governed by a Board of Trustees. The Education Review Office regularly reviews the administration, management, teaching and learning of schools, and makes recommendations to maintain standards (Knight, 2000). Possible negative impacts of high-stakes assessment are limited in New Zealand, whereby a variety of optional, nationally-validated assessment approaches are available to measure students' progress in relation to the national curriculum (Brown, 2004; Crooks, 2011; Nusche et al., 2012).

Assessment in New Zealand has a broad focus on improving teaching and learning, and less focus on summative testing, according to international standards (Crooks, 2011). There are currently no imposed nationwide tests for accountability at the primary level, although recommended national standards for each age group have been identified alongside curriculum documents - student achievement is measured against these (Nusche et al., 2012).

It could be asserted that these standards may form the beginnings of further education reform and higher accountability measures. Based on the aforementioned themes within the literature, the potential detrimental effects on student learning would need to be heavily considered and negated if New Zealand educational accountability measures increase.

Future Implications

It is essential to consider the implications of imposed standardised testing for students as assessment can provide useful and necessary information about student learning, and can also have a significant impact on a young person's academic life and development. Literature regarding assessment for accountability strongly supports two key themes among the implications for primary student learning; negative effects, and the unreliability of large-scale test scores as reflective measures of student ability.

The New Zealand educational accountability model is considered unique and 'low-stakes' according to international standards. Future New Zealand education policy must heavily consider the aforementioned findings and clarify the purpose of education, in order to encourage effective learning experiences and outcomes for students. There is an overall lack of readily available research in the area of consequential student learning outcomes at the primary level. There is also minimal research regarding teachers' perceptions of educational assessment, and the impact of content in beginning-teacher programmes and resultant effects on views of assessment and accountability.

Teachers play an essential role in the education system, therefore their views should be considered. Future research must consider how the aforementioned negative effects of assessment could be mitigated in an accountable system, and whether accountability is necessary. Teachers will need to be prepared to navigate accountability and performance pressures from conflicting paradigms in education, and may need to keep parents informed and involved regarding educational decisions.

References

- [1] Biesta, G. (2010). What is Education For? *Good Education in an Age of Measurement*. London: Paradigm Publishers, p.10-27.
- [2] Brown, G. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education*, 11(3), 301-318.
- [3] Crooks, T. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, 58(4), 438-481
- [4] Crooks, T. (2011). Assessment for learning in the accountability era: New Zealand. *Studies in Educational Evaluation*, 37(1), 71-77.
- [5] Gariepy, K., Spencer, B., & Couture, J. (Eds) (2009). *Educational Accountability. Professional Voices from the Field*. Sense Publishers. The Netherlands.
- [6] Herman, J. (2005). *CSE Report 649: Making Accountability Work to Improve Student Learning*. CRESST/University of California, Los Angeles.
- [7] Knight, N. (2000). Assessment for Accountability, Reporting, and Improving Learning: Have we got the Balance Right? *Delta*, 52(2), 39-62.
- [8] Marlow, R., Norwich, B., Ukoumunne, O., Hansford, L., Sharkey, S., & Ford, T. (2014). A comparison of teacher assessment (APP) with standardised tests in primary literacy and numeracy (WIAT-II). *Assessment in Education: Principles, Policy and Practice*, 21(4), 412-426.
- [9] Ministry of Education (1994). *Assessment: Policy to Practice*. Wellington: Learning Media.
- [10] Nusche, D., Laveault, D., MacBeath, J., & Santiago, P. (2012). *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Publishing. <http://dx.doi.org/10.1787/9789264116917-en>
- [11] Podgursky, M., & Springer, M. (2007). Teacher Performance Pay: A Review. *Journal of Policy Analysis and Management*, 26(4), 909-949.
- [12] Putwain, D. (2008). Deconstructing test anxiety. *Emotional and Behavioural Difficulties*, 13(2), 141-155.
- [13] Sadler, R. (2000). A Bird's Eye View of Assessment in Schools. National Regional Seminars, Auckland.
- [14] Shutz, P., & Pekrun, R. (2011). *Emotion in Education*. Academic Press: Munich